Consistent Force Field Studies of Intermolecular Forces in Hydrogen-Bonded Crystals. 2. A Benchmark for the Objective Comparison of Alternative Force Fields

A. T. Hagler,* S. Lifson,* and P. Dauber

Contribution from the Chemical Physics Department, Weizmann Institute of Science, Rehovot, Israel. *Received August 3, 1978*

Abstract: A benchmark for the objective comparison of intermolecular force fields of carboxylic acids and amides is established. It includes all experimental data used by most authors of empirical energy studies hitherto. In order to demonstrate the utility of such a benchmark, a detailed comparison between three force fields is performed. First, root mean square deviations of various significant properties for the set of acid crystals and for that of amides are presented for all three force fields. This is followed by an individual comparison of these deviations in all the 14 carboxylic acids and the 12 amides of the benchmark. Many observations and conclusions of interest are derived from the benchmark comparison. Among them, the following are of a more general character: (1) The 9-6-1 CFF potential gives a best overall fit to experiment. (2) The calculated amide crystal properties fit generally better than those of carboxylic acid in all force fields, probably because the amide group forms twice as many hydrogen bonds, and therefore tends as a rule to form more extended networks of hydrogen bonds in the crystals. (3) There is a significant correlation of the degree of fit of most individual crystals in all force fields. The "bad" crystals are generally bad in all force fields, and similarly for the "good" ones. Possible reasons for these trends are discussed.

Introduction

In the preceding paper,[1] we reported on the derivation of a force field for intermolecular interactions of carboxylic acids, based on the previously derived force field for amides.[2] The strategy adopted was to examine the transferability of the energy functions from amides to carboxylic acids, by seeking the smallest number of additions to the amide force field which would account for both carboxylic acids and amides.

Here we wish to establish and apply a "benchmark" which may be used for an objective evaluation of the various potential functions put forward in the literature. The need for a comparative evaluation of different force fields by objective criteria is deeply felt for a number of reasons. The first is that those who would like to apply one of the available sets of energy functions should be able to judge which is best, what sort of deviations in calculated structures and energetics to expect on their systems, and what sacrifice need be made, if any, in using a computationally inexpensive potential rather than a more "elaborate" one. As noted in the first paper,[1] analytical potential functions are being applied to a wide range of problems, from protein folding and dynamics to the effect of crystal forces on molecular conformation, yet no extensive comparative investigation such as that suggested here has been made. The second reason, which is related to the first, is that for the most part potential functions derived in the literature, even when for the same functional group, tend to use different systems for their data base, different methods for the derivation, and different criteria to evaluate the validity of the resultant functions. Thus, even were one prepared to do a comparative analysis, the information is not available in the literature and in fact additional calculations would need to be carried out.

In addition to the rationale given above, such a comparison is also needed to further investigate the physical meaningfulness of various terms and approximations in the different force fields. For example, in this paper we shall examine such questions as the basis for explicit hydrogen bond functions proposed to represent this interaction, and the transferability of potential parameters from the carbonyl oxygen to the hydroxyl oxygen. This is accomplished by comparing the success of the various functions in fitting the properties of the systems comprising the benchmark with the degree of success achieved by an optimized simpler form not containing the extra functions or parameters, on the same data base. Finally such a comparison helps to distinguish between artifacts of a particular force field and deviations common to various force fields. The latter pinpoint problematic systems, analysis of which may be used in order to further improve the analytical representation and understanding of the intermolecular forces.

For the present we compare three analytical representations suggested for amides and carboxylic acids. These include the two alternative force fields, 9-6-1 and 12-6-1, proposed by us for both amides[2] and carboxylic acids,[1] and the force field proposed by Scheraga and co-workers (MCMS)[3a,b] for amides and carboxylic acids. Since the publication of this force field Scheraga and co-workers have put forward a new model potential function, EPEN, to describe inter- and intramolecular interaction energies.[3c] The parametrization for amides and acids has just recently appeared[3d] and it would be of interest to extend the benchmark to include this, as well as the force fields derived recently by Smit[4] for carboxylic acids, and by Caillet and Claverie[5] who use different criteria from the others for both amides and acids. There is one category of analytical representation conspicuously absent from this list. There are no entries representing force fields derived by fitting ab initio energy surfaces. This is becoming an increasingly popular method for deriving energy functions,[6-8] because of the ease with which one may obtain the "experimental" data. It would be of great interest to apply such functions to the proposed benchmark, but unfortunately at present the published data is insufficient.

Method

Construction of Benchmark. Data Base. An important prerequisite for an objective and reliable evaluation of potential functions is the availability of an extensive data base of high-quality experimental results. We have included for the purposes of the present study all experimental data, involving crystal structures, sublimation energies, and several carboxylic acid dimerization energies and structures, used by us and Momany et al.[3] for the derivation of the energy functions. Additional crystals, not used in the optimization of the force fields, were also included as a further test of the functions. All data used in the derivation of each force field was included in order to avoid giving any force field an undue advantage. Altogether the data base comprises 14 carboxylic acid crystals, 12 amide crystals, and 2 carboxylic acid gas phase dimers. We have centered the data base around crystal properties because these are among the most accurate and extensive data available (the

0002-7863/79/1501-5122$01.00/0 © 1979 American Chemical Society

Table I. Root Mean Square Deviations of Properties Calculated for Carboxylic Acids and Amides by the Various Force Fields

property	units	no. of terms	rms dev 12-6-1	9-6-1	MCMS
		Acids			
energy	kcal/mol	12	2.486	2.053	2.118
UCV length	Å	42	0.489	0.307	0.604
UCV angle	deg	17	3.456	2.856	4.465
volume	Å^3	14	15.911	16.772	18.876
$d < 4$	Å	14	0.247	0.190	0.322
H···O dist	Å	16	0.062	0.072	0.058
O···O dist	Å	16	0.047	0.071	0.041
C-O···O angle	deg	16	11.071	9.881	14.048
O···O=C angle	deg	16	7.843	7.760	11.786
H···O=C angle	deg	16	12.362	12.144	17.985
180° − O-H···O	deg	16	8.491	7.732	11.710
		Amides			
energy	kcal/mol	6	1.574	1.930	8.446
UCV length	Å	36	0.208	0.235	0.261
UCV angle	deg	14	1.824	1.261	2.385
volume	Å^3	12	7.057	17.797	13.951
$d < 4$	Å	12	0.145	0.145	0.164
H···O dist	Å	30	0.049	0.059	0.056
N···O dist	Å	30	0.055	0.055	0.076
C-N···O angle	deg	22	3.337	3.575	4.071
N···O=C angle	deg	22	5.931	5.502	9.257
H···O=C angle	deg	30	5.830	5.609	7.329
180° − NH···O	deg	30	4.396	3.894	4.093

same reason they are used extensively to derive potential functions as noted in ref 1-4 and references cited therein).

Calculation of Crystal Structures and Energies. The procedure for calculating the minimum energy and structure is given in part 3.[9] It differs from most procedures in that symmetry is not assumed but derived, with the only constraint being the observed number of molecules per unit cell. In general the derivation of symmetry is trivial if one starts at the experimental structure, and does not serve as a test of potential functions.[2b,9,10] (It *is* important to relax symmetry constraints when treating hypothetical crystal structures, where the initial structure may not constitute a viable alternative[10].) Nevertheless, one case was found,[9] butyric acid (out of the 14 acids and 12 amides minimized), where the experimentally observed symmetry was lost upon minimization with all functions tried.[9] This is discussed in more detail in the following paper. The quadratically convergent minimization algorithm VA09A taken from the Harwell subroutine library was used to carry out the minimization. Convergence was based on the achievement of derivatives smaller than 10^{-5} kcal/mol·Å or 10^{-5} kcal/mol·rad. This is an important consideration as the minimization procedures used by previous authors may not have been convergent, either because only a subset of variables was minimized or nonconvergent minimization procedures such as steepest descent were used.

For the purposes of the benchmark the root mean square deviations of the unit cell vectors from their experimental values for each crystal are calculated as well as various other energetic and structural properties for each potential. This allows us to pinpoint particular crystals with large deviations and to compare various functions, which are the purposes of this study. Individual cell parameters are not examined and for the most part no attempt is made here to investigate the sources of the deviations at the molecular level. This is carried out in the following paper for the acids.[9]

Results and Discussion

A summary of the root mean square deviations for all observables constituting the data base is presented in Table I for the three force fields discussed above (12-6-1, 9-6-1, and MCMS). Here the deviations for each molecule within a family have been averaged, and no distinction is made between different molecules, although the values for different kinds of observables are given separately and we have listed separately the results for amides and acids. A total of 439 observables are represented in this table.

This is the most useful way for presenting the data in order to assess and compare the validity of different force fields. The properties included in this table represent the energy, as given by sublimation energy of the crystals, structure as represented by unit cell vector lengths (UCV lengths), angles (UCV angles), unit cell volumes, and the root mean square deviations between observed distances less than 4 Å in the crystal and their corresponding calculated values.[2] The latter is perhaps one of the best measures of differences in crystal structures. Large deviations in the usual measures, unit cell vector lengths, angles, and translation and rotation of the asymmetric unit, often tend to compensate for one another leading to a small change in energy and only moderate changes in the overall relative orientations of the molecules.[2] Finally we include several geometric properties of the hydrogen bond because of the importance of and interest in this particular interaction. These include two distances, H···O and O···O (or N···O in the case of amides), and four angles characterizing the main features of this interaction.

Several general observations as well as the relative merits of the different force fields and answers to the questions posed above concerning the assumptions in these force fields emerge from consideration of Table I. The most pertinent question is of course which force field is "best"? From the results presented in Table I it would appear that overall the 9-6-1 force field is superior to the other two. This is mainly an outcome of its clear superiority in the acids, especially in the important properties including lattice energy, UCV length, and interatomic distances less than 4 Å. The situation is not completely black and white, however, as the 12-6-1 gives a slightly better fit to the amides. In general, we advise the application of more than one ("reasonably good") potential function to a given problem where possible, to assure that the result is not an artifact of a given potential.[10-12]

Transferability of Oxygen Parameters and Use of Explicit H-Bond Potential. At this pont it is worth noting that both the 12-6-1 and 9-6-1 potentials for amides and acids involved the optimization of a total of ten energy parameters used in the simple functional form of eq 1 of the previous paper (excluding hydrocarbon parameters which were transferred).[1,2] These along with assumptions concerning combining rules and electroneutrality[1,2] defined the force field. The MCMS potential involved the optimization of a total of 12 parameters for the same functional groups, including those for an explicit hydrogen bond potential of the form

$$V_{HB} = A/r^{12}_{O\cdots H} - B/r^{10}_{O\cdots H} \qquad (1)$$

This term is included in addition to the terms in the 12-6-1 potential where it replaces the 12-6 (Lennard-Jones) term for the H···O interactions. These parameters along with the use of partial charges from CNDO/2, dispersion coefficients from the Slater-Kirkwood equation, and an assumed combining rule for mixed interactions completely define the MCMS potential.[3] The additional two parameters (12 vs. 10) do not improve the fit over the simpler functional form and in many properties the fit is significantly worse. Thus it may be concluded that neither the use of an explicit potential to represent the hydrogen bond nor separate parameters for the hydroxyl oxygen atom (i.e., not transferring the carbonyl oxygen parameters to the hydroxyl oxygen) improve the representation of the energy surface within the context of this potential. This will be discussed further below.

Table II. Lattice Energies of Amide and Carboxylic Acid Crystals (kcal/mol)[a]

molecule	exptl	9-6-1 calcd	9-6-1 dif	12-6-1 calcd	12-6-1 dif	MCMS calcd	MCMS dif
			Acids				
formic acid	−15.2	−13.3	1.9	−14.6	0.6	−13.7	1.5
acetic acid	−16.3	−15.2	1.1	−16.0	0.3	−15.3	1.0
propionic acid	−17.7	−17.6	**0.1**	−17.5	**−0.2**	−16.2	1.5
butyric acid	−19.2	−19.1	**0.1**	−18.9	0.3	−18.1	1.1
valeric acid	−20.2	−21.3	−1.1	−20.7	−0.5	−19.7	0.5
α-oxalic acid	−24.8	−27.4	−2.6	−29.0	−4.2	−27.5	−2.7
β-oxalic acid	−23.5	−26.8	*−3.3*	−28.6	*−5.1*	−28.7	*−5.1*
malonic acid		−28.6		−29.6		−28.2	
methylmalonic acid		−29.6		−30.0		−27.2	
succinic acid	−29.3	−32.0	−2.7	−32.9	−3.60	−30.5	−1.2
glutaric acid	−29.0	−31.1	−2.1	−32.0	−3.00	−29.7	−0.7
adipic acid	−32.1	−34.5	−2.3	−34.0	−1.90	−31.9	**−0.2**
suberic acid	−35.4	−37.5	−2.1	−37.1	−1.70	−34.7	−0.7
sebacic acid	−39.6	−41.9	−2.3	−41.1	−1.50	−42.8	−3.2
			Amides				
oxamide	−28.2	−25.4	*2.8*	−27.5	0.7	−19.7	8.5
malonamide	−28.8	−31.0	−2.7	−31.7	*−2.9*	−19.2	9.6
succinamide	−32.3	−34.3	−2.0	−33.5	−1.2	−22.1	*10.2*
glutaramide		−36.9		−36.0		−21.7	
adipamide		−38.8		−38.3		−25.4	
urea	−22.2	−23.4	−1.2	−23.0	**−0.8**		
formamide	−17.5	−15.7	1.7	−16.5	1.0	−10.9	**6.6**
diketopiperazine	−26.0	−27.0	**−1.0**	−27.7	−1.7	−19.4	**6.6**
LL-dimethyldiketopiperazine		−27.4		−27.4		−19.5	
cyclopropanecarboxamide		−23.7		−21.4		−13.5	
N-methylacetamide	(∼−18)	−15.7	(∼2)	−16.8	(∼1)	−11.4	(∼6)
suberamide		−44.1		−42.5		−28.2	

[a] In this and subsequent tables, we use boldface type to single out the cases of the best fit in each force field and italics to indicate the largest deviations.

Table III. Root Mean Square Deviations of the Lengths (Å) and Angles (deg) of the Unit Cell Vectors

molecule	9-6-1 lengths	9-6-1 angles	12-6-1 lengths	12-6-1 angles	MCMS lengths	MCMS angles	σ[a]
			Acids				
formic acid	0.22	0.0	0.16	0.0	0.50	0.0	0.25
acetic acid	0.62	0.0	0.97	0.0	1.24	0.0	0.18
propionic acid	0.09	0.5	0.17	0.3	0.28	2.0	
butyric acid	0.43	*5.8*	0.25	5.3	0.27	7.0	0.07
valeric acid	0.18	0.7	0.20	1.4	0.18	0.6	
α-oxalic acid	*0.70*	0.0	*1.11*	0.0	*1.39*	0.0	0.67
β-oxalic acid	0.23	1.9	0.83	*8.1*	0.84	*9.7*	
malonic acid	0.15	2.9	**0.13**	3.2	0.14	4.6	
methylmalonic acid	0.14	1.6	0.20	1.7	0.21	2.2	
succinic acid	0.12	0.3	0.18	1.2	**0.13**	1.4	0.12
glutaric acid	0.20	0.4	0.32	1.7	0.39	0.7	
adipic acid	**0.09**	0.4	0.30	1.1	0.43	2.3	
suberic acid	0.11	**0.1**	0.15	2.5	0.20	3.0	0.12
sebacic acid	**0.09**	0.9	0.19	**0.2**	0.20	**0.1**	0.13
			Amides				
oxamide	0.03	1.1	0.16	*3.2*	0.10	*4.4*	0.09
malonamide	0.10	0.3	0.11	0.8	0.27	0.2	
succinamide	**0.05**	1.0	0.07	**0.1**	0.11	2.5	0.09
glutaramide	*0.49*	*2.5*	*0.47*	2.5	0.44	1.8	
adipamide	0.12	**0.2**	**0.06**	1.0	0.05	0.7	0.05
urea	0.09	0.0	0.07	0.0			
formamide	0.40	1.2	0.36	2.1	*0.45*	0.5	0.13
diketopiperazine	0.09	1.4	0.10	0.2	0.11	2.5	
LL-dimethyldiketopiperazine	0.27	1.4	0.21	1.1	0.24	1.2	
cyclopropanecarboxamide	0.29	1.0	0.10	0.5	0.31	**0.1**	
N-methylacetamide	0.18	0.0	0.20	0.0	0.30	0.0	0.05
suberamide	0.17	0.8	0.09	0.6	**0.03**	1.0	

[a] Root mean square deviations of lengths and angles reported by MCMS (see Discussion).

Amides vs. Acids. Consideration of the root mean square deviations of the amides as compared to the acids leads to the conclusion that all three force fields are more successful in fitting the amide structures. The only exception to this is the large root mean square deviation between the calculated and observed amide sublimation energies in the case of the MCMS

Table IV. Root Mean Square Deviations of Short (<4 Å) Interatomic Distances (Å)

molecule	9-6-1	12-6-1	MCMS
	Acids		
formic acid	0.13	0.12	0.23
acetic acid	0.15	0.35	0.40
propionic acid	0.16	0.15	0.21
butyric acid	*0.34*	0.29	0.39
valeric acid	0.14	0.12	**0.11**
α-oxalic acid	0.28	*0.38*	*0.51*
β-oxalic acid	0.15	0.33	0.40
malonic acid	0.20	0.23	0.35
methylmalonic acid	0.21	0.21	0.27
succinic acid	**0.08**	**0.11**	0.15
glutaric acid	0.12	0.15	0.28
adipic acid	0.19	0.27	0.36
suberic acid	0.18	0.25	0.31
sebacic acid	0.16	0.23	0.29
	Amides		
oxamide	0.12	0.19	0.16
malonamide	0.06	0.08	0.15
succinamide	**0.05**	0.09	0.09
glutaramide	*0.22*	*0.23*	0.21
adipamide	0.06	0.05	**0.06**
urea	0.09	0.08	
formamide	0.21	0.18	0.23
diketopiperazine	0.13	0.14	0.14
LL-dimethyldiketopiperazine	0.21	0.16	0.18
cyclopropanecarboxamide	0.19	0.13	0.13
N-methylacetamide	0.15	0.21	*0.25*
suberamide	0.06	**0.03**	0.07

force field. It is worthwhile making a short digression here to discuss this apparently anomalous result as it sheds some light on why the additional parameters in this force field do not improve the fit to experimental properties. Comparison of the energy contributions with the 12-6-1 potential shows that the

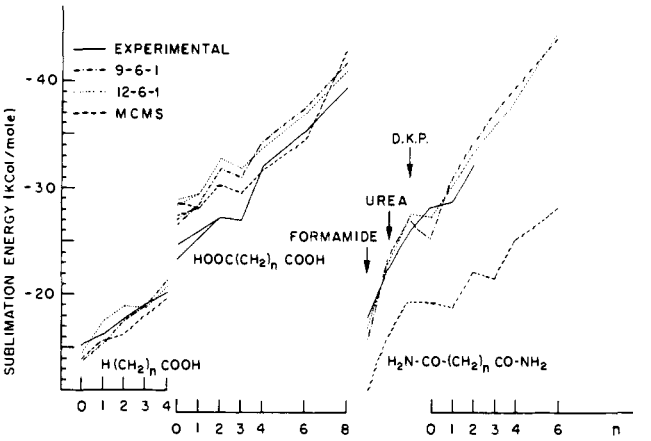

Figure 1. Sublimation energies of monocarboxylic acids ($H(CH_2)_nCOOH$, n = 0, 1, 2, 3, 4), dicarboxylic acids ($COOH(CH_2)_nCOOH$, n = 0, 1, 2, 3, 4, 6, 8), diamides ($H_2NCO(CH_2)_nCONH_2$, n = 0, 1, 2, 3, 4, 6), diketopiperazine (DKP), urea, and formamide (experimental values (——) and those calculated by 9-6-1 (-----), 12-6-1 (····) and MCMS (– – –) potentials).

Lennard-Jones (12-6) contributions are approximately the same in the two force fields but that the electrostatic contribution in the MCMS is much too small. The small Coulomb contribution arises from the use of partial charges obtained from population analysis of CNDO/2 molecular orbitals, and the use of a dielectric constant of 2 rather than 1. It has been pointed out that the use of partial atomic charges obtained from molecular orbital calculations in conformational energy calculations is of doubtful validity (ref 12 and references cited therein). In the case of the acids the MCMS potential reproduces the sublimation energy fairly well, since the large acid hydrogen-bond contribution compensates for the small electrostatic contribution. Thus it would appear that the additional parameters are used in this force field to compensate to some extent for the small electrostatic term.

Table V. Unit Cell Volumes (Å^3)

molecule	exptl	9-6-1 calcd	9-6-1 dif	12-6-1 calcd	12-6-1 dif	MCMS calcd	MCMS dif
			Acids				
formic acid	199	209	10	197	**−2**	189	−10
acetic acid	302	303	**1**	294	−8	292	−10
propionic acid	402	387	−15	389	−13	395	**−7**
butyric acid	515	480	*−35*	491	−24	492	−23
valeric acid	595	571	−24	588	−7	586	−9
α-oxalic acid	313	332	19	302	−11	286	*−27*
β-oxalic acid	157	166	9	150	−7	135	−22
malonic acid	211	216	5	199	−12	188	−23
methylmalonic acid	273	260	−13	249	−24	246	*−27*
succinic acid	251	252	**1**	241	−10	234	−17
glutaric acid	628	625	−3	597	*−31*	603	−25
adipic acid	355	343	−12	342	−13	341	−14
suberic acid	456	439	−17	441	−15	439	−17
sebacic acid	553	529	−24	535	−18	535	−18
			Amides				
oxamide	87.9	90.8	2.9	90.5	2.6	84.5	−3.4
malonamide	950	956	6	948	−2	914	*−35*
succinamide	534	528	−6	543	9	531	−3
glutaramide	676	653	−23	660	*−16*	659	−16
adipamide	353	353	**0**	359	6	355	2
urea	151	148	−3	148	−3		
formamide	230	226	−4	228	−2	211	−19
diketopiperazine	238	247	9	251	13	245	7
LL-dimethyldiketopiperazine	180	177	−3	180	**0**	181	**1**
cyclopropanecarboxamide	933	885	*−48*	933	**0**	926	−8
N-methylacetamide	453	453	**0**	457	3	456	3
suberamide	931	903	−8	933	2	920	−10

Table VI. Hydrogen Bond Geometry in Acid Crystals. The H···O Distance (Å)

molecule	exptl	9-6-1 calcd	9-6-1 dif	12-6-1 calcd	12-6-1 dif	MCMS calcd	MCMS dif
formic acid	1.58	1.71	*0.13*	1.61	0.03	1.68	*0.10*
acetic acid	1.65	1.69	0.04	1.58	−0.07	1.67	0.02
propionic acid	1.63	1.71	0.08	1.62	−0.01	1.68	0.05
butyric acid	1.62	1.71	0.09	1.63	0.01	1.68	0.06
valeric acid	1.65	1.71	0.06	1.62	−0.03	1.68	0.03
α-oxalic acid	1.80	1.75	−0.05	1.61	*−0.19*	1.67	−0.13
β-oxalic acid	1.67	1.73	0.06	1.64	−0.03	1.68	0.01
malonic acid	1.67	1.72	0.05	1.65	−0.02	1.69	0.02
	1.71	1.78	0.07	1.71	**0.00**	1.68	−0.03
methylmalonic acid	1.60	1.70	0.10	1.60	**0.00**	1.68	0.08
	1.74	1.73	**−0.01**	1.64	−0.10	1.68	−0.06
succinic acid	1.61	1.65	0.04	1.56	−0.05	1.67	0.06
glutaric acid	1.68	1.71	0.03	1.62	−0.06	1.68	**0.00**
adipic acid	1.62	1.71	0.09	1.61	−0.01	1.68	0.06
suberic acid	1.65	1.72	0.07	1.62	−0.03	1.68	0.03
sebacic acid	1.64	1.72	0.08	1.63	−0.01	1.68	0.04

Table VII. Hydrogen Bond Geometry in Amide Crystals. The H···O Distance (Å)

molecule	exptl	9-6-1 calcd	9-6-1 dif	12-6-1 calcd	12-6-1 dif	MCMS calcd	MCMS dif
oxamide	2.02	2.00	−0.02	1.96	−0.06	1.95	−0.07
	1.94	1.95	0.01	1.96	0.02	1.96	0.02
malonamide	1.90	1.94	0.04	1.98	0.08	1.95	0.05
	1.97	1.98	0.01	1.96	−0.01	1.94	−0.03
	2.31	2.40	0.09	2.35	0.04	2.14	*−0.17*
	1.95	1.92	−0.03	1.90	−0.05	1.95	**0.00**
	1.96	2.02	0.06	2.03	0.07	2.04	0.08
	1.97	2.02	0.05	2.01	0.04	2.00	0.03
	2.13	2.13	**0.00**	2.13	**0.00**	2.03	−0.10
	1.99	2.00	0.01	1.99	**0.00**	1.96	−0.03
succinamide	1.94	1.94	**0.00**	1.96	0.02	1.97	0.03
	1.95	1.92	−0.03	1.94	−0.01	1.98	0.03
glutaramide	2.02	1.98	−0.04	1.99	−0.03	2.04	0.02
	1.98	1.95	−0.03	1.97	−0.01	1.98	**0.00**
adipamide	1.98	1.98	**0.00**	1.97	−0.01	1.98	**0.00**
	2.06	2.01	−0.05	2.03	−0.03	1.97	−0.09
urea	2.07	1.96	*−0.11*	1.97	−0.10		
	2.09	2.15	0.06	2.12	0.03		
formamide	1.91	1.91	**0.00**	1.91	**0.00**	1.94	0.03
	1.95	1.99	0.04	1.95	**0.00**	1.95	**0.00**
diketopiperazine	1.85	1.96	*0.11*	1.94	0.09	1.93	0.08
LL-dimethyl-	1.91	2.01	0.10	2.02	*0.11*	1.96	0.05
diketopiperazine	1.93	2.04	*0.11*	1.99	0.06	1.98	0.05
cyclopropanecarbox-	1.99	1.95	−0.04	1.95	−0.04	1.97	−0.02
amide	1.92	1.87	−0.05	1.87	−0.05	1.92	**0.00**
	1.98	1.87	*−0.11*	1.92	−0.06	1.96	−0.02
	2.03	1.98	−0.05	2.04	0.01	2.03	**0.00**
N-methylacetamide	1.81	1.89	0.08	1.88	0.07	1.90	0.09
suberamide	1.96	1.91	−0.05	1.93	−0.03	1.95	−0.01
	1.98	1.93	−0.05	1.96	−0.02	1.96	−0.02

Returning to the comparison of the amides and acids, with the exception of the MCMS sublimation energy, the force fields fit the energy, UCV lengths, and short distances ("*d*" less than 4 Å) significantly better in the amides than the acids. One possible explanation for this observation is that in the amides all amide parameters were optimized while in the carboxylic acids all but one were transferred, without change from the amides. However, in the MCMS force field the amides and acids were treated on an equal basis[3] and the results of several other optimizations in the case of the 12-6-1 and 9-6-1 potentials[1] both indicate that this is not the source of the difference. Rather it would appear that the better fit of the amide properties is linked to the relative number of hydrogen bonds formed in the two classes of crystals. As can be seen from the list of the individual crystals included (given in Table II and subsequent tables), most of the amides included in the various studies are primary amides. Since a primary amide forms two hydrogen bonds per amide group (corresponding to the two amide protons), while the carboxylic acids form only one hydrogen bond per carboxyl group, the network of hydrogen bonds is twice as dense in the amides. It may be noted in Table I that in general the hydrogen bond distances are fit significantly better on the whole (~0.05 Å in both amides and acids) than all interatomic distances in the crystal (root mean square deviation = 0.3-0.6 in acids and ~0.15 in amides). Thus it would appear that the additional hydrogen bonds serve as additional "constraints" in the amides, leading to a better calculated structure. It follows from this that the deviation of the calculated properties of secondary amides would resemble the acids more than the primary amides, and this remains a subject for further study.

Individual Crystals. It is now worth turning to a summary

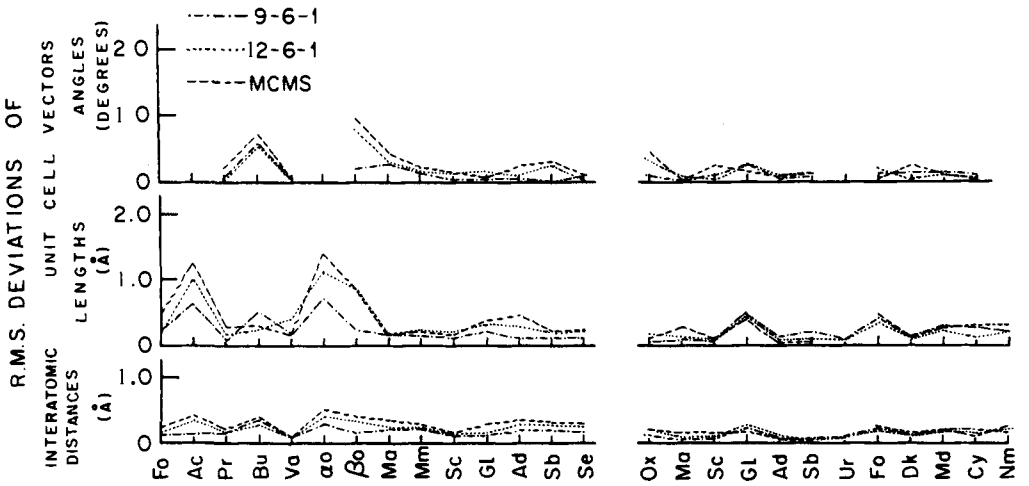

Figure 2. The root mean square (rms) deviations of structural properties of acids and amides. These properties include interatomic distances less than 4 Å, unit cell vector lengths, and angles. Abbreviations for acids: formic (Fo), acetic (Ac), propionic (Pr), butyric (Bu), valeric (Va), α-oxalic (αo), β-oxalic (βo), malonic (Ma), methylmalonic (Mm), succinic (Sc), glutaric (Gl), adipic (Ad), suberic (Sb), and sebacic (Se). For amides: oxamide (Ox), malonamide (Ma), succinamide (Sc), glutaramide (Gl), adipamide (Ad), suberamide (Sb), urea (Ur), formamide (Fo), diketopiperazine (Dk), dimethyldiketopiperazine (Md), cyclopropanecarboxamide (Cy), and *N*-methylacetamide (Nm).

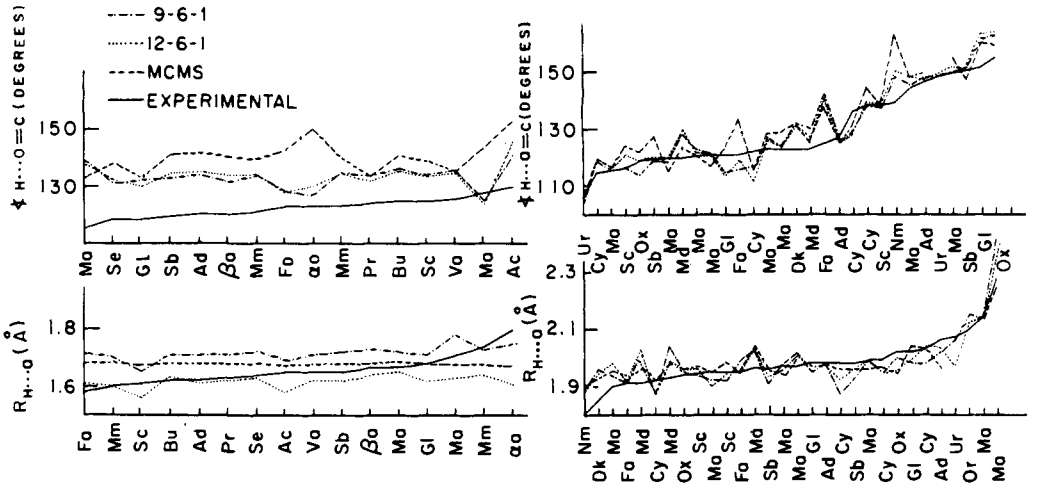

Figure 3. The hydrogen bond geometry (H···O distance and H···O=C angle) in acids and amides. Abbreviations for compounds as in Figure 2. Molecules were ordered according to increasing observed value of the property.

Table VIII. Hydrogen Bond Geometry in Acid Crystals. The H···O=C Angle (deg)

		9-6-1		12-6-1		MCMS	
molecule	exptl	calcd	dif	calcd	dif	calcd	dif
formic acid	122.5	127.7	5.2	127.8	5.3	142.6	20.1
acetic acid	130.2	141.5	11.3	144.9	14.7	153.6	*23.4*
propionic acid	124.0	133.5	9.5	132.2	8.2	134.2	10.2
butyric acid	125.0	136.3	11.3	135.5	10.5	141.4	16.4
valeric acid	125.7	136.2	10.5	134.8	9.1	135.0	**9.3**
α-oxalic acid	122.7	126.6	3.9	130.5	7.8	140.9	18.2
β-oxalic acid	120.3	132.3	12.0	133.5	13.2	140.8	20.5
malonic acid	127.5	124.9	**−2.6**	124.2	**−3.3**	143.8	16.3
	114.5	138.6	*24.1*	137.6	*23.1*	132.6	18.1
methylmalonic acid	121.1	134.3	13.2	134.2	13.1	140.4	19.3
	123.5	135.2	11.7	134.3	10.8	139.9	16.4
succinic acid	125.0	133.6	8.6	134.0	9.0	139.1	14.1
glutaric acid	118.4	132.1	13.7	129.9	11.5	133.6	15.2
adipic acid	120.2	134.0	13.8	135.2	15.0	141.7	21.5
suberic acid	119.1	133.1	14.0	134.4	15.3	141.0	21.9
sebacic acid	117.8	130.9	13.1	132.0	14.2	138.0	20.2

of the properties of the individual crystals, in order to determine if the average deviations are more or less representative, or whether there are large variations with, e.g., several crystals responsible for the major contribution to the total root mean square deviations. It is also of interest to see whether the pattern of deviations is common to the various functions or whether different analytical representations vary widely in which crystals are fit well and which badly.

Table IX. Hydrogen Bond Geometry in Amide Crystals. The H···O=C Angle (deg)

molecule	exptl	9-6-1 calcd	9-6-1 dif	12-6-1 calcd	12-6-1 dif	MCMS calcd	MCMS dif
oxamide	155	162	7	163	7	159	4
	119	114	−5	119	**0**	122	3
malonamide	121	123	2	122	1	121	**0**
	123	127	4	126	3	128	5
	149	150	1	151	2	155	5
	121	121	**0**	121	**0**	117	−4
	116	116	**0**	117	1	116	**0**
	120	119	−1	120	**0**	115	−5
	144	145	1	148	4	148	4
	123	124	1	124	1	129	6
succinamide	138	137	−1	138	0	138	**0**
	117	117	**0**	121	−4	124	7
glutaramide	151	161	10	162	11	160	9
	121	115	−6	114	−7	123	2
adipamide	127	127	**0**	127	**0**	126	−1
	146	148	2	147	1	149	3
urea	106	102	−4	102	−4		
	148	149	1	149	1		
formamide	125	142	*17*	141	*16*	138	13
	121	116	−5	119	−2	133	12
diketopiperazine	123	132	9	131	8	132	9
LL-dimethyl-diketopiperazine	120	129	9	130	10	123	3
	123	130	7	127	4	126	3
cyclopropanecarbox-amide	136	130	−6	128	−8	131	−5
	138	138	**0**	139	1	144	6
	115	119	4	116	1	120	5
	122	117	−5	112	−10	116	−6
N-methylacetamide	139	148	9	150	11	162	*23*
suberamide	119	119	**0**	120	1	127	8
	150	151	1	151	1	147	−3

Table X. Hydrogen Bond Geometry in Acid Crystals. The 180° − O-H···O Angle (deg)

molecule	exptl	9-6-1 calcd	9-6-1 dif	12-6-1 calcd	12-6-1 dif	MCMS calcd	MCMS dif
formic acid	6.8	7.4	**0.6**	7.6	**0.8**	20.0	13.2
acetic acid	16.7	9.8	−6.9	7.7	−9.0	10.8	−5.9
propionic acid	2.4	9.6	7.2	9.6	7.2	12.7	10.3
butyric acid	0.3	12.2	11.9	10.2	9.9	17.0	16.7
valeric acid	12.4	5.8	−6.6	7.6	−4.8	10.0	−2.4
α-oxalic acid	32.5	15.0	*17.5*	10.5	*−22.0*	13.2	*−19.3*
β-oxalic acid	5.2	9.4	4.2	9.5	4.3	16.8	11.6
malonic acid	10.8	6.9	−3.9	12.9	2.1	28.5	17.7
	8.4	18.3	9.9	20.4	12.0	19.7	11.3
methylmalonic acid	3.8	4.8	1.0	4.9	1.1	20.8	17.0
	10.5	18.3	7.8	14.6	4.1	10.5	**0.0**
succinic acid	0.0	9.3	9.3	9.9	9.9	14.9	14.9
glutaric acid	12.1	4.6	−7.5	2.5	−9.6	8.6	−3.5
adipic acid	6.1	10.9	4.8	9.5	3.4	16.4	10.3
suberic acid	10.2	9.1	−1.1	6.9	−3.3	13.8	3.6
sebacic acid	10.1	7.7	−2.4	5.9	−4.2	12.2	2.1

Sublimation Energies. The deviations in sublimation energies are given in Table II and in Figure 1. We can immediately identify the sublimation energies of the two oxalic acid crystals as problematic observables, which are the worst in all potentials. This deviation is discussed in more detail in the following paper.[3]

A second feature common to the three potentials is the negative deviation of the dicarboxylic acids. This is in contrast to the monocarboxylic acids where the calculated sublimation energy is in general less negative than observed. This may indicate that to some extent the energetic contribution of the carboxyl group is too negative, perhaps compensating for another term in the force field. Another possibility is that the dicarboxylic acids form intramolecular hydrogen bonds in the gas phase, thus lowering their sublimation energies. This would, however, require an anti-planar O=C—OH conformation which is 2-4 kcal less stable than the syn-planar structure, so it is not clear how much energy would be gained by such a hydrogen bond.[13] This pattern is a subject for further investigation, and may constitute a stringent test for further improvement of the potentials. Finally one of the most outstanding features in sublimation energies is the large deviation found for amides in the case of the MCMS. This was reflected in the total root mean square deviation and commented on above. Aside from the latter, the trends in the deviations are rather similar for the different potentials. This is remarkable especially considering the large contribution of the 10-12 hydrogen bond term in the MCMS potential, for which there is no corresponding explicit term in the 12-6-1 or 9-6-1 potentials.

Structural Properties. The root mean square deviations for the structural properties are given in Tables III-V and in

Table XI. Hydrogen Bond Geometry in Amide Crystals. The 180° − N-H···O Angle (deg)

molecule	exptl	9-6-1		12-6-1		MCMS	
		calcd	dif	calcd	dif	calcd	dif
oxamide	29	21	−8	20	−9	24	−5
	4	2	−2	4	**0**	7	3
malonamide	2	3	1	4	2	3	1
	13	10	−3	11	−2	12	−1
	41	42	1	44	3	43	2
	22	22	**0**	23	1	30	8
	4	4	**0**	3	−1	2	−2
	20	19	−1	19	−1	24	4
	28	27	−1	29	1	30	2
	20	21	1	20	**0**	25	5
succinamide	3	4	1	5	2	2	−1
	4	3	−1	5	1	5	1
glutaramide	28	18	*−10*	17	−11	20	−8
	11	11	**0**	14	3	12	1
adipamide	18	19	1	18	**0**	21	3
	35	33	−2	34	−1	32	−3
urea	12	16	4	17	5		
	28	28	**0**	28	**0**		
formamide	19	13	−6	16	−3	15	−4
	13	9	−4	4	−9	15	2
diketopiperazine	4	13	9	16	*12*	16	*12*
LL-dimethyl-diketopiperazine	11	18	7	19	8	10	−1
	13	15	2	12	−1	16	3
cyclopropanecarbox-amide	19	17	−2	22	3	19	**0**
	10	12	2	9	−1	4	−6
	10	4	−6	6	−4	8	−2
	18	18	**0**	18	**0**	18	**0**
N-methylacetamide	6	3	−3	3	−3	8	2
suberamide	6	7	1	7	1	9	3
	29	29	**0**	29	**0**	32	3

Figure 2. Again comparison of the root mean square deviations for the individual crystals as calculated with the individual potentials, particularly as represented in Tables III and IV, reveals that there are "problematic" systems and overall similarities of the goodness of fit among the different potential functions. Acetic, butyric, and again the two oxalic acids are seen to have the largest deviations in the acids, while glutaramide and formamide present the challenges in the amides. The deviations found with the 9-6-1 and 12-6-1 potentials in the acid crystals are discussed at the molecular level in the following paper, while the amide crystals were discussed previously. The deviations are smallest for the most part for the 9-6-1 and largest for the MCMS potentials as reflected by the overall root mean square deviation given in Table I, but, with the exception of β-oxalic acid, those crystals which are found to have large calculated deviations are "bad" in all potentials considered. Thus, despite the differences in the potentials, they seem to have some common deficiencies which are manifested in these crystals.

The last column in Table III, σ, represents the root mean square deviation of both the unit cell vector lengths and angles, as reported by MCMS,[3] and should be equal to the appropriately combined root mean square deviation of the two preceding columns. The apparently lower root mean square deviation obtained by MCMS arises because they minimized the energy only with respect to the unit cell vectors and angles. The rotational and translational degrees of freedom were kept fixed at their experimental values. Thus, as seen from the table, when starting from the experimental structure, the constrained minimum is much closer to the observed than the true minimum. The latter is clearly a more stringent test of potential functions and should be applied in order to avoid misleading, "apparently good fits" to experiment.

The Hydrogen Bond. Finally it is of interest to compare the deviations in hydrogen bond geometries for the individual crystals. These values are given in Tables VI-XI and in Figure 3. The usefulness of comparison of many systems is again brought out in the consideration of the hydrogen bonds. One is immediately struck by the constancy of the calculated value of the O···H distance in the acids (1.68 ± 0.01 Å) in the MCMS potential. This is not reflected in either the experimental values or in the calculated values of the other two potentials. Thus, although the root mean square deviation in O···H distance in acids is approximately the same for the three potentials (~0.06), it arises in a completely different way in the MCMS potential and the other two. The effect of the various packing modes is not felt in the MCMS potential owing to the strong O···H-O hydrogen bond potential (~6 kcal) which imposes an O···H length independent of environment. The different packing environments are reflected in the 9-6-1 and 12-6-1 potentials where the span in calculated distances is only slightly less than the observed variation. Here the root mean square deviation given in Table I represents a random error. The constant value of the O···H distance in the MCMS is related to the discussion of sublimation energies and partial charges given above. There it was noted that the large 10-12 contribution in acids was needed to compensate for the small electrostatic contribution. This O···H potential then results in an artificially strong constraint on the O···H distance. In the amides (Table VII), this behavior is not observed in the MCMS potential since the 10-12 O···H term is much weaker ($V_{min} \simeq 1$ kcal).

The situation with the angles is very different, but here again the potentials exhibit the same trend in deviations. For example, as seen in Table VIII, there is a strong bias for the calculated H···O=C angle to be larger than the experimental value, and this bias is observed in all three potentials. This bias also occurs in the amides, although not quite so dramatically. It was noted previously[2b] and suggested that it may arise because of the omission of an explicit representation of the lone-pair electrons in these potentials. The angular dependence of the hydrogen bond has been discussed in some detail, and

with these potentials the minimum energy for the *isolated* three atoms H···O=C occurs at 180°.[14] The angular dependence is very shallow, and most of the angular dependence of the hydrogen bond has been attributed to the constraints imposed by the other atoms in the molecule.[3,14] Explicit inclusion of the lone-pair orbitals might well change this angular dependence, however. Furthermore, it should be noted that the large deviations in structure often correspond to rather small energy differences.[2b,9] Thus even small change in the energetics of the angular dependence may result in large structural changes. This bias in the H···O=C angles would then seem to also provide an area which can both suggest and serve as a test of further improvements to the potential functions.

Conclusion

One of the purposes of this work was to attempt to provide some indication of the "errors" to be expected in applying these potentials to other systems. It is difficult to extrapolate, since the potentials were derived from crystal systems and these are in many cases *relatively* simple. However, when applying such potentials to other systems, for example, proteins, one should not expect to get a better representation of the structure than some of the *worst* deviations obtained here. Thus one should expect in general root mean square deviations in interatomic distances of the order of 0.3–0.4 Å due to the potentials alone (i.e., other complicating factors such as solvent effects will make this worse). As noted above, the hydrogen bonds are fit better than this, and one can reasonably assume that O···O or N···O distances are within 0.1 Å of the experimental ones, with larger deviations for the H···O=C angle of up to 30° not unreasonable. It follows that, if the "density" of hydrogen bonds is large, one can expect the overall deviations to be correspondingly smaller. The intermolecular energies should be good to ~10%.

It should be emphasized that these potential functions take into account all atoms including hydrogens and the deviations refer to *inter*molecular distances where there is no constraint on the movement of individual molecules. It is usual when considering larger systems to make simplifying assumptions, varying from considering the methyl group as a united atom to considering entire amino acid side chains as single atoms as well as assuming various parts of the molecule rigid. The above estimates clearly do not apply in such cases.

Acknowledgment. This research was partially supported by grants from the U.S.-Israel Binational Science Foundation (BSF), Jerusalem, Israel, and by the Israeli Academy of Sciences.

References and Notes

(1) Part 1: S. Lifson, A. T. Hagler, and P. Dauber, *J. Am. Chem. Soc.*, preceding paper in this issue.
(2) (a) A. T. Hagler, E. Huler, and S. Lifson, *J. Am. Chem. Soc.*, **96**, 5319 (1974); (b) A. T. Hagler and S. Lifson, *ibid.*, **96**, 5327 (1974).
(3) (a) F. A. Momany, L. M. Carruthers, R. F. McGuire, and H. A. Scheraga, *J. Phys. Chem.*, **78**, 1595 (1974); (b) L. G. Dunfield, A. W. Burgess, and H. A. Scheraga, *ibid.*, **82**, 2609 (1978); (c) L. L. Shipman, A. W. Burgess, and H. A. Scheraga, *Proc. Natl. Acad. Sci. U.S.A.*, **72**, 543 (1975); (d) J. Snir, R. A. Nemenoff, and H. A. Scheraga, *J. Phys. Chem.*, **82**, 2527 (1978).
(4) P. H. Smit, Ph.D. Dissertation, Department of Structural Chemistry, Rijksuniversiteit, The Netherlands, 1978.
(5) J. Caillet and P. Claverie, *Biopolymers*, **13**, 601 (1974).
(6) (a) E. Clementi, F. Cavallone, and R. Scordamaglia, *J. Am. Chem. Soc.*, **99**, 5531 (1977); (b) R. Scordamaglia, F. Cavallone, and E. Clementi, *ibid.*, **99**, 5545 (1977); (c) G. Boles and E. Clementi, *ibid.*, **99**, 5550 (1977).
(7) J. C. Owicki and H. A. Scheraga, *J. Am. Chem. Soc.*, **99**, 7413 (1977).
(8) S. Swaminathan, R. J. Whitehead, E. Guth, and D. L. Beveridge, *J. Am. Chem. Soc.*, **99**, 7317 (1977).
(9) Part 3: A. T. Hagler, P. Dauber, and S. Lifson, *J. Am. Chem. Soc.*, following paper in this issue.
(10) A. T. Hagler and L. Leiserowitz, *J. Am. Chem. Soc.*, **100**, 5879 (1978).
(11) D. A. Brant, *Annu. Rev. Biophys. Bioeng.*, 1, 369 (1972).
(12) (a) A. T. Hagler, L. Leiserowitz, and M. Tuval, *J. Am. Chem. Soc.*, **98**, 4600 (1976); (b) J. Bernstein and A. T. Hagler, *ibid.*, **100**, 673 (1978).
(13) L. Leiserowitz, *Acta Crystallogr., Sect. B*, **32**, 775 (1976).
(14) A. T. Hagler, E. Huler, and S. Lifson, "Peptides, Polypeptides and Proteins", E. R. Blout, F. A. Bovey, M. Goodman, and N. Lotan, Eds., Wiley, New York, 1974, p 35.

0002-7863/79/1501-5130$01.00/0 © 1979 American Chemical Society